Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 18, No. 2, 2026

Research on Multidimensional Evaluation of College Foreign Language Teacher Training

Yali Feng

School of Foreign Languages, Hubei University of Technology, Wuhan, Hubei, China

Abstract: Against the backdrop of global communication and the digital transformation in education, the professional competency of college foreign language teachers plays a crucial role in shaping both the quality of foreign language instruction and the success of cross-cultural interactions. Currently, foreign language teacher training faces challenges such as uniform content, a singular assessment framework, and inadequate adaptability for developing language skills. These issues hinder the ability to address the disruptions introduced by emerging technologies like ChatGPT and to fulfill the contemporary goal of effectively "telling China's stories". This study examined the current state of foreign language teacher training across Chinese colleges and universities and formulated a comprehensive evaluation framework that encompassed various subjects, dimensions, indicators, and methodologies. Additionally, it proposed enhancement strategies such as skill-specific training, optimization of hybrid models, and the establishment of an evaluation-feedback loop. The findings aim to offer theoretical insights and practical guidance for effectively implementing college foreign language teacher training programs while promoting systematic advancements in the quality of foreign language education.

Keywords: College Foreign Language Teachers; Teacher Training; Multidimensional Evaluation; Professional Development; Quality of Foreign Language Education.

1. Introduction

As a fundamental element in nurturing international talent and facilitating cross-cultural communication within higher education institutions, enhancing the quality of foreign language education hinges on the ongoing advancement of teachers' professional skills. In today's globalized context, the necessity of effectively "telling China's stories" imposes heightened expectations on college foreign language educators regarding their abilities in cross-cultural communication and international engagement. Following the pandemic, hybrid teaching—combining online and offline methods—has become standard practice. The challenges related to technical utilization and classroom interaction management that emerged during emergency remote teaching underscore the critical need for teacher training[1]. Moreover, the emergence of artificial intelligence tools like ChatGPT has significantly influenced traditional foreign language pedagogy, prompting educators to redefine their roles and concentrate on developing essential competencies that AI cannot replicate, such as critical thinking and cultural interpretation[2]. Consequently, advancing the training and development of foreign language teachers is vital for reforming college-level foreign language education.

Nonetheless, several issues persist within current programs aimed at training college foreign language teachers. The content often lacks specificity and fails to meet the diverse developmental needs associated with five essential language skills (listening, speaking, reading, writing, and translating) [3]. Meanwhile, evaluations primarily assess outcomes without providing feedback on processes or considering multidimensional adaptability. This gap results in a disconnection between training outcomes and actual teacher requirements. Furthermore, many educators possess a vague understanding of their personal development needs and tend to engage in various trainings without achieving targeted enhancements in their professional skills. In light of these challenges, establishing a scientifically sound multidimensional evaluation framework along with targeted strategies for training improvement has become essential to address current shortcomings and foster the professional growth of college foreign language instructors.

Current research on evaluating foreign language teacher training tends to concentrate on singular dimensions or broad indicators while neglecting the unique aspects of language skill development. Thus, existing evaluation systems lack precision and depth. This study adopts a multidimensional evaluation perspective by incorporating various evaluators—including teachers, organizers, students, and experts—to create an evaluation framework tailored to the specifics of foreign language instruction. This approach not only enriches theoretical perspectives surrounding teacher training evaluation but also offers fresh insights into teacher development research within foreign language education. Additionally, examining how evaluation interacts with training deepens the educational principle of "enhancing training through evaluation", thereby adding practical relevance to theories concerning teacher professional development.

The proposed strategies for training enhancement outlined in this study can serve as actionable references for policymakers and designers involved in college-level foreign language teacher education. These strategies aim to address issues such as uniformity in training content and superficial evaluations while offering feasible implementation pathways and assessment tools across different levels of colleges and universities. In doing so, the study intends to ultimately enhance both relevance and effectiveness in teacher training initiatives. It also provides concrete examples and clear guidance for individual foreign language educators seeking to accurately identify areas needing improvement amidst a complex educational landscape while achieving sustainable career progression.

2. Literature Review

2.1. Strategies for Foreign Language Teacher Training

One group of recent research has focused on practice orientation and has transitioned from model innovation to scenario-based adaptation with particular attention to the skill adaptability of college foreign language teachers. Some studies have explored innovative strategies supported by emerging technologies, which has centered around the combined use of artificial intelligence and virtual reality. By creating scenarios such as international academic conferences and cross-cultural classroom interactions, teachers were able to enhance their speaking skills and abilities in managing cultural conflicts through immersive experiences[2]. Other studies have dwelled on skill/group-specific strategy segmentation. Strategies tailored to meet various skill demands have been progressively emerging. For instance, Gou[4]introduced a three-phase strategy involving "machine translation assistance–manual proofreading–case review", which aimed at enhancing translation skills. Additionally, Brunsmeier[5]proposed a research-led training strategy, which included collaborative efforts in applying for foreign language education projects and publishing papers on teaching methodologies, effectively addressing issues related to low motivation for participation among senior educators.

Another group of studies have closely aligned with policy initiatives like the National Training Program and the Provincial Training Program. They have increasingly concentrated on resolving prevalent challenges faced by college foreign language teachers, such as gaps between training and actual teaching practices or resource shortages, resulting in a research framework characterized by model innovation – mechanism support – targeted adaptation for diverse groups[6][7][8]. Some other studies have delved into research support on mechanisms ensuring sustainable training[9]. To combat issues surrounding insufficient sustainability within training programs, scholars have begun prioritizing mechanism development initiatives [10].

2.2. Assessment of Foreign Language Teacher Training Effectiveness

Previous studies on assessing the effects of foreign language teacher training have targeted evaluation indicators and dimensions. First, tailored evaluation instruments were created to cater to the unique aspects of foreign language training. Sadoughi and Hejazi[11] designed the Foreign Language Teacher Support Scale to measure four types of perceived teacher support, namely, emotional, instrumental, appraisal, and informational support. In addition, technical innovation in evaluation methods were introduced. For example, Shi and Wang[12] developed a system for analyzing classroom teaching behaviors through AI video analysis that monitors factors such as teachers' interaction frequency during speaking skill teaching and time spent explaining cross-cultural cases. By integrating this data with students' engagement metrics, a process evaluation model was established, achieving an impressive prediction accuracy regarding the effectiveness of speaking skill training, significantly surpassing the traditional final exam's accuracy rate. Another study delved into the evaluation of teacher education in Turkey[13]. Notably, English teachers and lecturers thought that the content knowledge, pedagogical content knowledge, and general culture courses in the teacher training program were insufficient. Furthermore, English teachers stated that they could not apply what they learned during their training to real-life conditions. Therefore, the foreign language teacher training program should be revised and relevant research is needed.

2.3. Research Gaps

The review of relevant literature indicates that both domestic and international investigations into college foreign language teacher training have examined the strategies for and assessment of foreign language teacher training. However, three critical structural issues remain unresolved.

First, there exists an absence of perspective on skill differentiation. Current literature often overlooks the distinct independence among the five core foreign language skills—listening, speaking, reading, writing—and translating when demonstrating necessity or designing strategies/evaluations. It fails to clarify why certain skills like translation exhibit higher demand or which strategies are suitable per specific skill types. Moreover, it lacks specialized indicators tailored per individual skills within evaluations, hence impeding practical support directed towards targeted skills-based training initiatives.

Second, the research into scenario adaptability is inadequate. Much existing work is largely centered around generalized university settings or English major contexts without acknowledging varying scenarios such as distinctions between science/engineering-focused institutions versus comprehensive universities or public/professional languages along with English compared against minority languages. Thus, gaps exist particularly concerning academic English needs pertinent among science/engineering university faculty alongside identified shortages experienced within cross-cultural resources available for minority-language instructors, which constrains broader applicability across diverse educational contexts.

Third, there exists a blank "training-evaluation" linkage mechanism that connects training strategies with evaluative feedback. Presently available research tends not only separate strategic designs from effect evaluations but also neglects reciprocal feedback loops whereby evaluative insights could refine strategic implementations. As a result, it ultimately fails to address pressing challenges surrounding uniformity found amidst similar trainings coupled alongside ambiguous outcome interpretations.

To tackle these issues head-on, this study proposes adopting "multidimensional evaluation" as an essential connective element focused upon differentiated demands inherent within each respective aspect associated amongst five core collegiate-level linguistic proficiencies. In so doing, it aims to construct an innovative multidimensional evaluative framework, which is organized according both situational variables along defined specialties. It intends to ultimately enable the formulation of targeted intervention strategies culminating closed loops wherein systematic evaluative processes duly inform subsequent adaptations. The study aims to provide theoretical insights alongside pragmatic pathways that can rectify contemporary organizational challenges for the current collegiate-level pedagogical frameworks surrounding foreign-language instructor preparations.

3. Methodology

3.1. Research Questions

Drawing from the research context and literature review, this study centers on three primary research questions:

a) What structural challenges are present in the content, format, and assessment of current training for college foreign language teachers? Are there any notable differences in training requirements among educators based on their years of experience, professional titles, and types of universities?

b) How can a comprehensive evaluation system for college foreign language teacher training be developed to meet the needs associated with the five foreign language skills? What should its core dimensions, indicators, and weights be?

c) Utilizing the multidimensional evaluation framework, what specific and actionable strategies should be established for enhancing training for college foreign language educators?

3.2. Participants

The study employed purposive sampling to select participants including foreign language instructors, decision-makers involved in training, and experts in foreign language education from 12 varied universities across Hubei Province. The sampled participants were designed to reflect diversity concerning university type, years of teaching experience, professional titles, and other relevant factors. The composition of the sampled participants is detailed as follows:

a) Distribution by university type: 4 comprehensive universities; 3 normal universities; 3 science and engineering universities; 2 private institutions.

b) Teacher sample (N=682): Gender breakdown – 523 females (76.7%) and 159 males (23.3%); languages instructed – 598 English teachers (87.7%), 32 Japanese teachers (4.7%), 21 French teachers (3.1%), and 31 teachers of other languages (4.5%); teaching roles – 496 public foreign language instructors (72.7%) versus 186 specialized foreign language educators (27.3%); years of teaching experience – 126 with less than three years (18.5%), 358 with three to ten years (52.5%), and 198 exceeding ten years (29%); professional designations – 132 teaching assistants (19.4%), 386 lecturers (56.6%), 145 associate professors (21.3%), and 19 full professors (2.8%).

c) Additional participants include: ten training decision-makers who manage teacher training within academic affairs offices or schools of foreign languages at various universities; five experts in foreign language education who are professors or doctoral supervisors at prominent domestic universities specializing in teacher development and applied linguistics.

3.3. Research Instruments

3.3.1. Questionnaire

Informed by Qu et al.[6] and Zhang's[14] evaluation scales for foreign language teacher training, this study developed a questionnaire addressing the current situation and needs pertaining to such training programs, differentiated into two versions: one for foreign language teachers and another for decision-makers overseeing training initiatives.

The trainee version encompassed four dimensions: participation in training activities (5 items), requirements regarding content and format of training programs (12 items), assessment of training effectiveness (10 items), awareness of a multidimensional evaluation system (8 items), totaling to an overall count of thirty-five items employing a Likert scale ranging from one to five (1=strongly disagree; 5=strongly agree), alongside two open-ended questions aimed at gathering specific feedback.

The decision-maker version comprised three dimensions: design aspects of the training plan (8 items), implementation status of current trainings(6 items), evaluation mechanisms used in trainings(7 items), culminating in twenty-one total items also utilizing a Likert scale with one open-ended question.

The development process involved several steps: initial item identification through literature reviews coupled with expert consultations; followed by revisions based on assessments from three experts in the field along with two management figures—leading to adjustments on six ambiguously phrased questions; ultimately concluding with a pre-survey involving eighty individuals that led to the removal of three poorly discriminating items through item analysis—resulting in a finalized formal questionnaire.

Reliability tests yielded Cronbach's α coefficients indicating strong internal consistency: .923 for the trainee version and .897 for the decision-maker version; additionally, KMO tests showed values suggesting suitability for factor analysis (.886 for trainees) along with significant results from Bartlett's test.

3.3.2. Semi-Structured Interview

To address the research questions effectively, two distinct sets of semi-structured interview guides were created—one tailored for foreign language educators while another catered to decision-makers/experts within foreign language education systems.

Key focus areas within the teacher interview guide encompassed experiences related to participation in trainings along with satisfaction levels, prioritization regarding essential skill areas needing attention during trainings, identification of key issues within existing programs, plus recommendations pertaining to establishing an effective multidimensional evaluation framework.

Conversely, discussions outlined within expert/decision-maker interviews involves critical aspects tied to designing effective plans for trainings while also addressing shortcomings seen within current evaluations—a consideration toward developing components such as dimensions/indicators necessary within a multifaceted evaluation system alongside potential integration methods linking both evaluations alongside trainings.

The interview outlines underwent refinement via two rounds featuring expert feedback ensuring both relevance as well as logical coherence throughout queries leading up towards final drafts comprising eight-to-ten central questions whilst maintaining flexibility through spaces allocated for follow-ups aimed at deeper explorations into participant insights.

3.3.3. Documents

Three different foreign language teacher training programs were selected for case studies, representing a range of university types. The selection criteria included a minimum program duration of over one year and well-defined objectives and evaluation frameworks that addressed at least three out of five essential linguistic competencies, with involvement from at least fifty educators.

Specifically, the first was a Digital Enhancement Program for Foreign Language Teaching Abilities, which represented a comprehensive institution. The second was a Specialized Training Initiative in Cross-Cultural Communication Instruction, which reflected standards found in traditional

educational settings. The third was an Academic English Reading/Writing Training Program, which showcased practices typical of science and engineering-oriented institutions. The training plans, summary reports, evaluation documents, and classroom observations of these three programs were collected.

4. Data Collection and Analysis

4.1. Data Collection Process

Between March and May 2024, a total of 800 questionnaires were distributed. Altogether, 750 were sent to trainees and 50 to decision-makers. Both were conducted online via Questionnaire Star. Out of these, 682 valid responses were received, resulting in an effective response rate of 85.25%. Specifically, there were 632 valid trainee responses with an effective recovery rate of 84.3%, while the decision-maker version achieved a full recovery rate of 100%.

From June to July 2024, interviews were conducted both online (via video) and in person. Each session lasted between 40 to 60 minutes, with all discussions recorded and transcribed into text format, culminating in approximately 150,000 words of interview material.

Information regarding the implementation processes and outcomes of training programs was gathered by reviewing training plans, summary reports, evaluation documents from participating universities, classroom observations, and teacher interviews. This resulted in the establishment of a comprehensive case analysis database.

4.2. Data Analysis

The data analysis commenced with quantitative methods. SPSS version 26.0 was utilized to evaluate the questionnaire data through various statistical techniques including descriptive statistics (frequency counts, means, standard deviations), difference tests (independent sample t-tests and one-way ANOVA), correlation assessments, and factor analyses to illustrate current training conditions as well as variations in demand and structural components within the evaluation system.

Subsequently, qualitative analysis was performed using Nvivo version 12 software to code and interpret the interview transcripts alongside case data. This process followed a structured approach, which included open coding and axial coding. The opening coding identified key concepts and preliminary categories. The axial coding organized logical relationships among these categories.

Lastly, triangulation validation was employed to merge both quantitative and qualitative findings to enhance the credibility and validity of the results. For instance, questionnaire data provided an overview of prevalent training issues while interview insights and case studies elucidated underlying causes for these challenges.

5. Results

5.1. Current Situation and Problems of College Foreign Language Teacher Training

5.1.1. Characteristics of Training Participation and Needs

In terms of participation rates and frequency, 82.3% of teachers engaged in various training programs over the past three years. However, notable discrepancies were observed in how often they participated. Teachers with fewer than three years of experience attended an average of 2.3 times per year, while those with 3 to 10 years participated 1.8 times, and teachers with more than 10 years had an average of 1.2 times. This trend illustrates that “the longer the teaching tenure, the lower the participation frequency” ($F=12.345$, $p<0.001$). Regarding professional titles, teaching assistants and lecturers had significantly higher participation rates compared to associate professors and professors ($F=8.762$, $p<0.01$).

When analyzing the demand for training content, descriptive statistics revealed that teachers prioritized skill training as follows: translation skills (mean 4.12) > writing skills (mean 3.98) > speaking skills (mean 3.85) > reading skills (mean 3.62) > listening skills (mean 3.45). Additionally, beyond skill training, there was a strong demand for digital teaching resource development (mean 4.05), AI technology-assisted teaching (mean 3.96), and cross-cultural teaching design (mean 3.88).

The independent sample t-test indicated significant differences in demand across various groups. Public foreign language teachers required more training in listening and speaking skills than their professional counterparts ($t=3.214$, $p<0.01$; $t=2.876$, $p<0.05$), whereas professional foreign language teachers showed a greater need for translation and writing skill training ($t=4.123$, $p<0.001$; $t=3.567$, $p<0.01$). Furthermore, faculty at science and engineering universities expressed a markedly higher demand for academic English reading, writing, and translation training compared to those at comprehensive or normal universities ($F=9.876$, $p<0.001$).

5.1.2. Structural Problems of Training Content and Form

Firstly, there was a significant issue of content uniformity. Altogether, 68.5% of educators felt that current training materials “lack relevance” (mean 3.87), while 45% reported “high levels of content repetition across different trainings” (mean 3.62). In interviews, many educators noted that “most training focuses on general instructional methods without targeting specific areas like listening or translating”, which provided limited practical assistance for public foreign language instruction (Teacher Interview No.T23). Additionally, comments were made regarding AI technology-related sessions being overly conceptual without integrating practical applications linked to language skills, rendering them ineffective post-training (Teacher Interview No.T156).

Secondly, the formats used for these trainings were often one-dimensional and rigid. In total, 72% of participants primarily experienced “offline lectures” as their mode of attendance (mean 3.95), with hybrid models comprising only 28%. However, online modules mostly included pre-recorded sessions lacking interactive components or personalized feedback mechanisms. Merely 15% engaged in hands-on formats such as workshops or scenario simulations, struggling to address real-world skill enhancement needs.

Thirdly, a disconnect from actual teaching practices persisted. Altogether, 63% believed that the training content was “too theoretical and fails to connect with classroom realities” (mean 3.78). Many interviewed teachers expressed difficulty implementing strategies taught during trainings within large classroom settings due to their impracticality in public foreign language contexts (Teacher Interview No.T345).

5.1.3. Existing Dilemmas of Training Evaluation System

The first dilemma was the narrow scope of evaluation subjects. Only 23% incorporated student evaluations into

their assessments. Meanwhile, just 12% involved expert evaluations. While a dominant majority (85%) relied on "training sign-in + final exam" as primary evaluative measures, all focused predominantly on organizers' assessments or self-evaluations by instructors.

The second was the lack of diverse evaluation criteria. In total, 67% reported that existing evaluations merely monitored participation rates without addressing obvious improvements in the outcomes of competency (mean 3.82). In addition, only 18 .3 % conducted follow-up assessments regarding teaching efficacy post-skill enhancement efforts. Evaluation metrics predominantly revolved around process indicators like "attendance" or "homework completion", failing to accurately gauge mastery levels or effects on instructional application.

Lastly, there existed an inadequate feedback system. A significant 78.9% indicated they had "never received detailed reports on training evaluation feedback" (mean 4.02) , while just over 9.8% noted that results were utilized for refining future trainings. Fewer than 5% established a complete cycle encompassing "training – evaluation – feedback – optimization".

5.2. Development of a Multidimensional Evaluation System for College Foreign Language Teacher Training

5.2.1. Selection and Refinement of Evaluation Indicators

The core indicators for the multidimensional evaluation system were established after three rounds of consultations with experts. This group comprised five specialists, including three professors and two associate professors, who collectively had an average teaching experience of 22.3 years across fields such as foreign language teacher development, applied linguistics, and educational evaluation. The expert authority coefficient was recorded at Cr=0.89, which satisfied the required standards. The positive coefficients from the expert consultations were 100%, 100%, and 80% respectively, while the coordination coefficients were noted as 0.72, 0.81, and 0.86 ($p<0.001$), indicating a significant level of consensus among the experts. Ultimately, four primary dimensions, fifteen secondary dimensions, and forty-two tertiary dimensions were selected (refer to Table 1).

5.2.2. Allocation of Weights to Evaluation Indicators

To assign weights to each indicator, the analytic hierarchy process was employed using Yaahp version 10.3 software for calculations. The consistency tests revealed that all judgment matrices exhibited CR values below 0.1, demonstrating a satisfactory level of consistency along with a rational distribution of weights. The weightings for the primary dimensions included teaching ability (0.35), professional literacy (0.25), training adaptability (0.20), cross-cultural literacy (0.12), and information literacy (0.08). Among the secondary dimensions, those with the highest weights included teaching effectiveness across five skills (0.18), alignment of training content (0.12), and enhancement of research capabilities (0.10). For tertiary dimensions, leading skill indicators were improvement in translation accuracy (0.042), enhancement of writing logic (0.038), and boosting oral expression confidence (0.035).

5.3. Comprehensive Structure of Multidimensional Evaluation System

A cohesive four-in-one framework for the multidimensional evaluation system was established encompassing "subject – dimension – indicator – method". Firstly, the evaluation system consisted of evaluation subjects, including trained teachers (weight: 0.3), training organizers (weight: 0.25), students (weight: 0.25), and foreign language education experts (weight: 0.2). This forms a collaborative multi-subject evaluation approach.

Secondly, the evaluation system included evaluation dimensions. In total, there were five areas, namely, teaching ability, professional literacy, information literacy, cross-cultural literacy, and training adaptability. Each was supported by corresponding second-level and third-level indicators (see Table 1).

Finally, the evaluation system comprised of evaluation methods. It integrated quantitative assessment methods such as questionnaires, skill tests, and teaching effect data analysis weighted at 0.5. Besides, it also incorporated qualitative assessment methods including classroom observations, interviews, and teaching reflection journals weighted at 0.3. The process evaluations mainly focused on training participation quality. The task completion weighted at 0.3, alongside summative evaluations like skill improvement tests and assessments of teaching effectiveness weighted at 0.2, achieving a comprehensive multi-faceted evaluation approach.

6. Discussion

Based on the developed multidimensional evaluation framework and the key issues identified during the investigation, four specific strategies have been recommended.

6.1. Targeted Skills Training Strategy for Language Teaching Adaptation

To address the varied developmental requirements across five foreign language skills, modular training programs tailored to each skill have been created. For listening/speaking training, emphasis is placed on cross-cultural oral communication, constructing and utilizing listening corpora, designing interactive oral classrooms, and incorporating real-life scenario simulations and immersive teaching techniques. Additionally, foreign experts are invited for specialized guidance. In reading/writing training, modules focus on academic English reading and writing, critical reading strategies, and writing process evaluations. This is complemented by high-quality textbook analyses and case studies of student work to enhance teachers' capabilities in text interpretation and instructional support.

The training for translation skills centers on essential areas like applying translation technology, bilingual conversion techniques, and translating culturally significant terms. Workshops focused on practical translation exercises and industry seminars aim to fulfill the translation demands inherent in international communication. Furthermore, a tiered training approach is implemented based on factors such as years of teaching experience, professional titles, and individual weaknesses. New teachers receive foundational skills and classroom management training. Younger teachers concentrate on research proficiency and integrating technology. Senior educators focus on pedagogical innovation and academic leadership.

Table 1. Indicators and Weights of Multidimensional Evaluation System for College Foreign Language Teacher Training

First-level Dimension (Weight)	Second-level Dimension (Weight)	Third-level Dimension (Weight)
Teaching Ability (0.35)	Listening Skill Teaching (0.06)	Adaptability of listening material selection (0.021); Diversity of listening training methods (0.019); Improvement effect of students' listening ability (0.020)
	Speaking Skill Teaching (0.07)	Rationality of oral interaction design (0.025); Accuracy of pronunciation guidance (0.020); Improvement of students' oral expression confidence (0.025)
	Reading Skill Teaching (0.06)	Effectiveness of reading strategy teaching (0.022); Depth of text interpretation (0.018); Improvement of students' reading speed and comprehension (0.020)
	Writing Skill Teaching (0.08)	Pertinence of writing task design (0.028); Detail of correction feedback (0.022); Optimization of students' writing logic and expression (0.030)
	Translation Skill Teaching (0.08)	Practicality of translation skill teaching (0.027); Effectiveness of bilingual conversion guidance (0.023); Improvement of students' translation accuracy (0.032)
Professional Literacy (0.25)	Language Knowledge Renewal (0.08)	Mastery of language theory (0.030); Innovation in vocabulary and grammar teaching (0.025); Reserve of disciplinary frontier knowledge (0.025)
	Research Ability Improvement (0.10)	Ability to apply for teaching and research projects (0.035); Level of academic paper writing (0.035); Application of teaching research methods (0.030)
	Textbook Interpretation and Development (0.07)	Grasp of key and difficult points in textbooks (0.028); Ability to develop school-based textbooks (0.022); Adaptability of textbooks to teaching goals (0.020)
Information Literacy (0.08)	Teaching Technology Application (0.04)	Proficiency in teaching platform operation (0.015); Ability to develop digital resources (0.015); Application of AI-assisted teaching tools (0.010)
	Technology Integration Innovation (0.04)	Integration of technology and skill teaching (0.020); Effectiveness of online-offline teaching connection (0.020)
Cross-Cultural Literacy (0.12)	Cross-Cultural Teaching Design (0.05)	Rationality of cultural element integration (0.025); Creation of cross-cultural scenarios (0.025)
	International Communication Ability (0.04)	Foreign language expression of Chinese culture (0.020); Cross-cultural conflict handling (0.020)
	Cultural Interpretation Ability (0.03)	Interpretation of cultural-loaded words (0.015); Comparative teaching of Chinese and foreign cultures (0.015)
Training Adaptability (0.20)	Content Matching (0.08)	Adaptability to skill needs (0.040); Adaptability to teaching practice (0.040)
	Form Applicability (0.06)	Diversity of training forms (0.030); Flexibility of learning time (0.030)
	Feedback Effectiveness (0.06)	Timeliness of feedback (0.025); Pertinence of feedback (0.025); Effect of feedback application (0.010)

6.2. Optimization Strategy for Hybrid Training Mode in Digital Education

In the post-pandemic landscape, hybrid training has emerged as a primary method for teacher development, necessitating enhancements in its implementation effectiveness. For online components, a digital platform can be established that consolidates high-quality MOOC resources, live lectures, virtual simulation training sessions, among others. Flexible learning schedules should be created to accommodate teachers' needs for intermittent study. AI-assisted tools like pronunciation evaluation systems and writing correction software could offer tailored feedback. For offline elements, activities such as specialized workshops, peer observation seminars, and face-to-face expert guidance should be organized to emphasize practical application and foster interactive dialogue. Collaborative training across institutions can facilitate resource sharing among educators from different universities. In an integrated online-offline format, theoretical knowledge would be acquired online while practical applications occur offline along with deep discussions facilitated through learning communities, creating a comprehensive learning cycle of "theory – practice – reflection – enhancement".

6.3. Evaluation-Driven Closed-Loop Training Strategy

Utilizing the multidimensional evaluation system as a foundation can help establish a connection between training initiatives and evaluation processes. Pre-training diagnostic assessments help accurately pinpoint teachers' skill gaps through testing or surveys that inform targeted training approaches. During the training phase, ongoing evaluations using methods such as classroom observations or performance assessments allow for real-time adjustments to content or delivery methods based on participant engagement.

Post-training evaluations provide a holistic view of effectiveness using various metrics including skill tests results along with feedback from students and expert reviews. Establishing a feedback mechanism allows detailed reports to be generated for both trainers and participants outlining strengths and areas needing improvement. This insight informs subsequent iterations of the training program, creating an iterative cycle of "evaluation–feedback–improvement–re-evaluation".

6.4. Support Strategy for Teacher Independent Development

The ultimate objective of these trainings is to foster educators' self-directed growth necessitating a robust support system. First of all, establishing professional development communities can encourage collaboration among foreign language instructors across different institutions through regular seminars or joint projects aimed at sharing successes. Secondly, offering personalized development resources tailored to individual assessment results—such as guidance in academic writing or project applications—can assist those with weaker research skills. Thirdly, implementing an incentive structure linking the success of these trainings with professional evaluations encourages motivation towards independent professional growth. Lastly, documenting exemplary cases from investigations into successful trainings will yield an "Excellent Training Case Collection", serving as a practical reference guide for educators seeking pathways toward their own development.

7. Conclusion

The evaluation system developed in this study addressed the shortcomings of current research that overlooked the unique aspects of language skills, thereby enhancing the theoretical framework surrounding foreign language teacher training assessment. The proposed closed-loop model of "evaluation – training – feedback" enriched the educational philosophy of "enhancing training through evaluation", offering a new practical framework for theories related to teacher professional development. This system supplied scientific evaluation instruments and targeted strategic guidance for decision-makers involved in college foreign language teacher training, aiding in the resolution of issues such as uniformity in training content and ineffective outcomes. It also presented actionable implementation strategies for foreign language teacher training across various academic institutions, thereby increasing both relevance and impact. Furthermore, it outlined clear pathways and illustrative examples for the independent growth of foreign language educators, supporting their sustainable career advancement.

Nonetheless, it is important to note that the research samples are primarily drawn from universities located in Hubei Province, indicating a need for improved regional representation. Additionally, the long-term efficacy of the evaluation system requires validation through subsequent studies. Future investigations should aim to broaden the sample size to include universities from different regions and categories while further refining evaluation metrics. Longitudinal studies spanning 3–5 years could be conducted to assess the sustained effectiveness of training approaches. Furthermore, integrating artificial intelligence technology could facilitate the development of intelligent evaluation and training platforms, enhancing both accuracy and efficiency in assessment and educational practices.

Acknowledgments

This article is a product of the Philosophy and Social Science Research Project of Hubei Provincial Department of Education in 2023, which is titled "Research on College Foreign Language Teacher Training and Improvement Strategies from the Perspective of Multidimensional Evaluation" (Project number is 23Y020).

References

[1] Latif Muhammad M M Abdel. (2023). Emergency Remote Teaching Experience and Reflective Beliefs of Foreign Language Teachers in Saudi Universities [J]. Higher Education Research & Development, 42(2): 678-692.

[2] Moreno Gimenez, A., & Parker, R. (2025).Immersive inclusivity: AI-enhanced virtual reality for language and cultures learning. Proceedings of the Lancaster University Education Conference. Vol. 02.

[3] Zhang, J., & Peng, Y. (2024). Unlocking teacher agency for professional learning through connection-building: Evidence from a professional development program on China's standards of English language ability. The Asia-Pacific Education Researcher, 33(6), 1515-1523.

[4] Brunsmeier, S. (2024). Teachers as researchers: Voices from experts. In Tools, techniques and strategies for reflective second & foreign language teacher education: Insights from contexts around the world (pp. 161-176). Berlin, Heidelberg: Springer Berlin Heidelberg.

[5] Gou, S. (2025). Research on the optimization path of machine translation from a cognitive perspective: The collaborative development of human translation and machine translation. The Frontiers of Society, Science and Technology, 7(7).

[6] Ou, Q., Wu, S., Jiang, Y., He, Z., & Luo, H. (2025). From training to practice: A multi-group analysis of factors influencing K-12 teachers' integration of Digital Educational Resources (DERs). Plos one, 20(12), e0338543.

[7] Yang, L., Garcia-Holgado, A., & Martinez-Abad, F. (2023). Digital competence of K-12 pre-service and in-service teachers in China: A systematic literature review. Asia pacific education review, 24(4), 679-693.

[8] Zhi, H., Ravindran, L., & Amini, M. (2025). A systematic review of the changes, challenges, and implications of China's teacher education policies (2012 to 2023). International Journal of Evaluation and Research in Education, 14(1), 207-221.

[9] Alcantud-Díaz, M., & LLoret-Catalá, C. (2023). Bridging the gap between teacher training and society. Sustainable development goals (SDGs) in English as a foreign language (EFL). Globalisation, Societies and Education, 1-12.

[10] Arslan, S., & Curle, S. (2024). Institutionalising English as a foreign language teachers for global sustainability: Perceptions of education for sustainable development in Turkey. International Journal of Educational Research, 125, 102353.

[11] Sadoughi, M., & Hejazi, S. Y. (2022). Development and validation of foreign language teacher support scale (FLTSS). Language Testing in Asia, 12(1), 30.

[12] Zhang, W. (2022). The role of technology-based education and teacher professional development in English as a foreign language classes. Frontiers in Psychology, 13, 910315.

[13] Shi, S., Gao, J., & Wang, W. (2022). Classroom teaching behavior analysis based on artificial intelligence. In Artificial intelligence in education and teaching assessment (pp. 25-36). Singapore: Springer Singapore.

[14] Çelik, S. N., & Memduhoglu, H. B. (2022). The Evaluation of the English Language Teacher Education Program in Turkey. International Journal of Educational Methodology, 8 (4), 833-851.